Concurrent Functional Linear Models

Pang Du

Department of Statistics
Virginia Tech

Predicting precipitation profiles from temperature curves

- Precipitation is much harder to predict than temperature.
- It comes in two main forms:
  - *Drizzle*: Large low pressure zones drop moisture over many hours or days.
  - *Storms*: Convective, short violent storms with a lot of precipitation in a hurry, and spatially localized.
- Precipitation tends to be seasonal; more in the spring and fall than in the summer and winter.
- Temperature effect: temperature drop in summer often comes with precipitation, snowfall in winter comes when temperature is a little below freezing.

A model

- We can assume that climate zone is important.
- We will predict log precipitation: logging stabilizes variance and eliminates the positivity constraint.
- We will use the difference $\text{TempRes}_{mg}(t)$ between a temperature profile and the mean for the climate zone as a functional covariate.
- We can extend the functional ANOVA model to

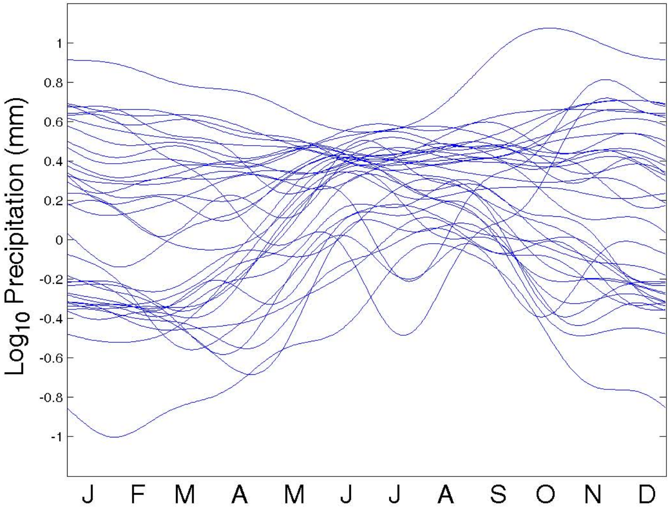
$$\log[\text{Prec}_{mg}(t)] = \mu(t) + \alpha_g(t) + \text{TempRes}_{mg}(t)\beta(t) + \epsilon_{mg}(t).$$

- We call this model *concurrent* because it assumes that the temperature today affects today's precipitation.

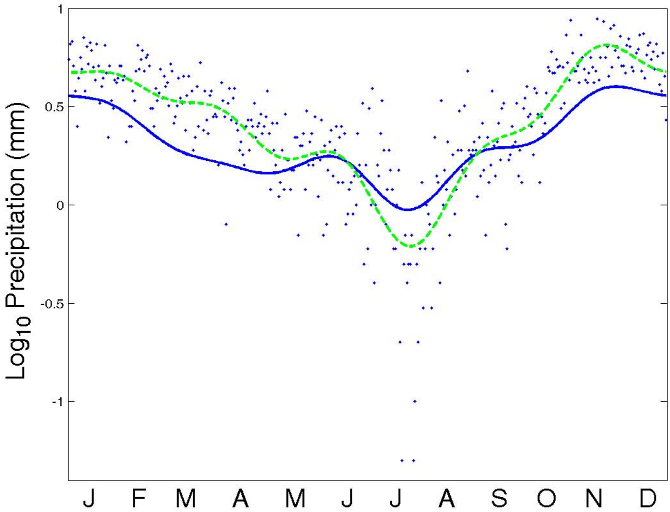
Preliminary steps

- Where precipitation was recorded as 0 mm, we changed it to 0.05 mm, half the minimum positive value.
- We used 11 Fourier series basis functions for precipitation with no roughness penalty.
- We used 21 Fourier series basis functions for temperature with no roughness penalty.

Log precipitation profiles



Log precipitation profile for Vancouver



Preliminary step highlights

- Prince Rupert is the rainiest place that averages nearly 12 mm of rain a day in October.
- The driest station is Resolute in the high arctic, where snowfall has a barely measurable rain equivalent of 0.1 mm per day in the winter.
- Vancouver: sharp drop in rainfall during the summer months, even including two days without precipitation in 34 years.

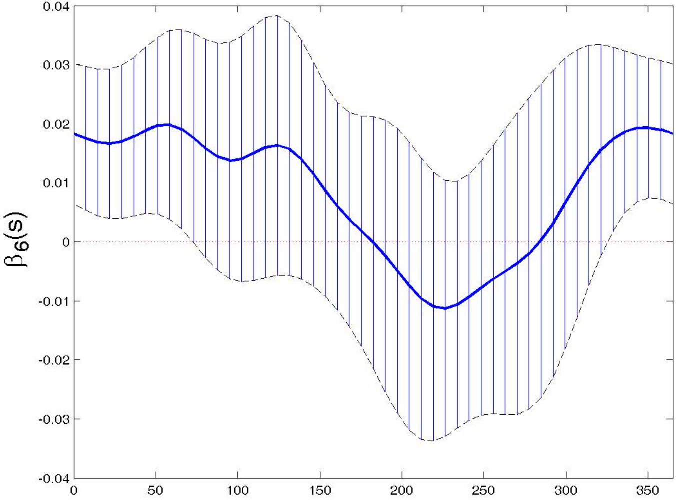
Fitting criterion and some results

- The fitting criterion was the unpenalized error sum of squares

$$\mathrm{LMSSE}(\mu, \alpha_g, \beta) = \sum_{g=1}^{4} \sum_{m=1}^{N_g} \int [\mathrm{LogPrec}_{mg}(t) - \mu(t) - \alpha_g(t) - \mathrm{TempRes}_{mg}(t)\beta(t)]^2 dt$$

- The resulting root-mean-squared-residual was 0.19 mm.
- When we dropped TempRes(t) from the model, this increased to 0.20 mm.
- The temperature residual function don't seem to improve the fit by much.

Estimated $\beta(t)$ and its point-wise CIs



More about temperature effect

- As we see in the following plot, the only place where temperature appears to make a contribution is the mid-winter, from December to February.
- However, the CIs are point-wise and don't provide sufficient conditions for making such a claim.
- A better way to investigate this is to construct a functional contrast (or probe).

A probe for the winter effect

- The confidence limits are point-wise; we need a measure of the temperature influence accumulated over the winter months.
- We could propose a contrast function or linear probe
$$\xi(t) = \cos[2\pi(t - 64.5)/365],$$
where the shift value of 64.5 is the empirical mark of mid-winter by finding the low point in the mean precipitation profile.
- Why this probe?
- $\xi(t)$ has its positive peak at the mid-winter and negative peak in the summer.

Probe results on the winter effect

- The inner product of the regression coefficient function with this probe
$$\int_0^{365} \cos[2\pi(t-64.5)/365]\beta(t)dt = 2.32$$
in effect accumulates information across the entire year about the difference between the summer and winter influence in temperature residual.
- The estimated standard error of this probe is 0.77, giving a t-ratio of 3.0.
- Conclusion: if the mean temperature residual for a weather station is high in winter (e.g., a marine station like Prince Rupert), then precipitation will also be high for that station relative to other stations within the same climate zone.

Model

- General form of a concurrent functional linear model:

$$y_i(t) = \sum_{j=1}^{q} z_{ij}(t)\beta_j(t) + \epsilon_i(t).$$

- or in matrix notation:

$$\mathbf{y}(t) = \mathbf{Z}(t)\boldsymbol{\beta}(t) + \boldsymbol{\epsilon}(t).$$

- The least squares criterion:

$$\text{LMSSE}(\boldsymbol{\beta}) = \int [\mathbf{y}(t) - \mathbf{Z}(t)\boldsymbol{\beta}(t)]'[\mathbf{y}(t) - \mathbf{Z}(t)\boldsymbol{\beta}(t)]dt.$$

Roughness penalties

- We may need a separate roughness penalty for each regression coefficient function:

$$\mathrm{PEN}_j(\beta_j) = \lambda_j \int [L_j \beta_j(t)]^2 dt.$$

- The penalized least squares criterion:

$$\mathrm{PLMSSE}(\boldsymbol{\beta}) = \int [\mathbf{y}(t) - \mathbf{Z}(t)\boldsymbol{\beta}(t)]'[\mathbf{y}(t) - \mathbf{Z}(t)\boldsymbol{\beta}(t)]dt + \sum_{j=1}^{q} \lambda_j \int [L_j \beta_j(t)]^2 dt.$$

Basis function expansions for $\beta_j(t)$

- Let regression function $\beta_j(t)$ have the expansion

$$\beta_j(t) = \mathbf{b}_j'\boldsymbol{\theta}_j(t)$$

  in terms of K_j basis functions $\theta_{jk}(t)$.
- For scalar predictors: the basis for their $\beta_j(t)$'s is the constant basis

$$\beta_{j1}(t) = 1.$$

Further notation

- Defining $K_\beta = \sum_{j=1}^{q} K_j$, we construct a vector **b** of length K_β as
$$\mathbf{b} = (\mathbf{b}_1', \ldots, \mathbf{b}_q')'.$$
- Now assemble q by K_β matrix function Θ as follows:
$$\Theta = \begin{bmatrix} \boldsymbol{\theta}_1' & 0 & \cdots & 0 \\ 0 & \boldsymbol{\theta}_2' & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{\theta}_q' \end{bmatrix}.$$
- We can now express our model as
$$\mathbf{y}(t) = \mathbf{Z}(t)\Theta(t)\mathbf{b} + \boldsymbol{\epsilon}(t).$$

Rewrite the penalties

- We also need to arrange the order K_j roughness penalty matrices

$$\lambda_j \mathbf{R}_j = \lambda_j \int [L\boldsymbol{\theta}_j(t)][L\boldsymbol{\theta}_j(t)]' dt$$

into the symmetric block diagonal matrix $\mathbf{R}$ of order K_β:

$$\mathbf{R} = \begin{bmatrix} \lambda_1 \mathbf{R}_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 \mathbf{R}_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_q \mathbf{R}_q \end{bmatrix}.$$

The normal equation

$$[\int \Theta'(t)\mathbf{Z}'(t)\mathbf{Z}(t)\Theta(t)dt + \mathbf{R}]\mathbf{b} = \int \Theta'(t)\mathbf{Z}'(t)\mathbf{y}(t)dt$$

- Numerical integrations on LHS: The scalar functions $w_{jl}(t) = \sum_{i=1}^{N} z_{ij}(t)z_{il}(t)$ play the role of *weighting functions* for the functional inner products

$$\int \boldsymbol{\theta}_j(t)\boldsymbol{\theta}_l'(t)w_{jl}(t), \quad j, l = 1, \ldots, q.$$

- Numerical integrations on RHS: The scalar functions $\sum_{i=1}^{N} z_{ij}(t)y_i(t)$ play the role of weighting functions for the functional inner products of the basis functions $\boldsymbol{\theta}_j$ with the unit function 1.

The confidence intervals

- Estimate the variance Σ_e of the smoothing residuals for a single response.
- Assuming independence of the observations, the variance of the whole response data matrix is

$$\text{Var}[\text{vec}(\mathbf{Y})] = \Sigma_e \otimes \mathbf{I}.$$

- Work out the linear mapping from the data to a functional probe $\rho(\beta_j)$ that is being estimated. Let us call this $\mathbf{M}_j$. Then

$$\text{Var}[\rho(\beta_j)] = \mathbf{M}_j'(\Sigma_e \otimes \mathbf{I})\mathbf{M}_j.$$

How do I work out mapping $\mathbf{M}_j$?

The probe $\rho(\beta_j)$ is three linear mappings from the data:

- Two mappings are the same as the functional-response-multivariate-covariate case:
  - The linear mapping from the raw data matrix $\mathbf{Y}$ to the coefficient matrix $\mathbf{C}$ defining the smoothed response functions $\mathbf{y}(t)$.
  - The linear mapping from the regression coefficient function coefficient vector $\mathbf{B}_j'$ to the value of the probe.
- We only need to work out the linear mapping from $\mathbf{C}$ to $\mathbf{B}_j'$.

The mapping from C to B

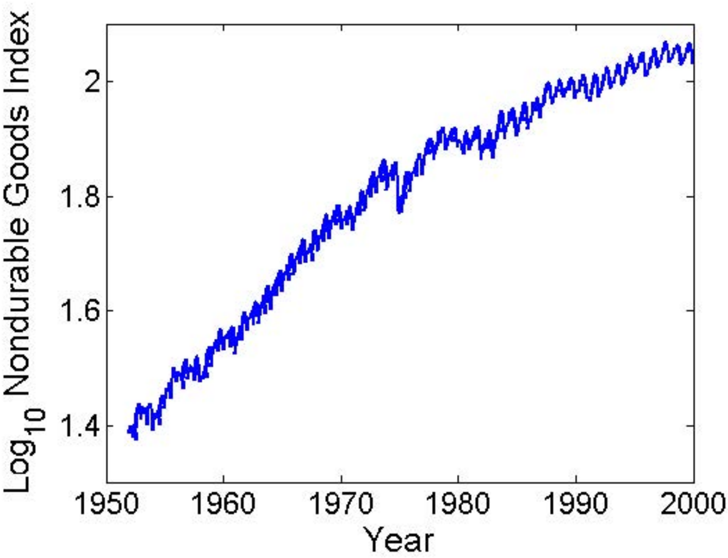
- Let ϕ be the length-K_y vector of basis functions for the response functions $\mathbf{y}(t)$ such that $\mathbf{y}(t) = \mathbf{C}\phi(t)$.
- From the normal equation:

$$\begin{aligned}\hat{\mathbf{b}} &= [\int \Theta'\mathbf{Z}'\mathbf{Z}\Theta + \mathbf{R}]^{-1}[\int \Theta'\mathbf{Z}'\mathbf{C}\phi] \\ &= [\int \Theta'\mathbf{Z}'\mathbf{Z}\Theta + \mathbf{R}]^{-1}[\int \phi' \otimes (\Theta'\mathbf{Z}')]\mathrm{vec}(\mathbf{C}).\end{aligned}$$

- Hence the mapping from $\mathbf{C}$ to $\mathbf{B}$ is

$$[\int \Theta'\mathbf{Z}'\mathbf{Z}\Theta + \mathbf{R}]^{-1}[\int \phi' \otimes (\Theta'\mathbf{Z}')]$$

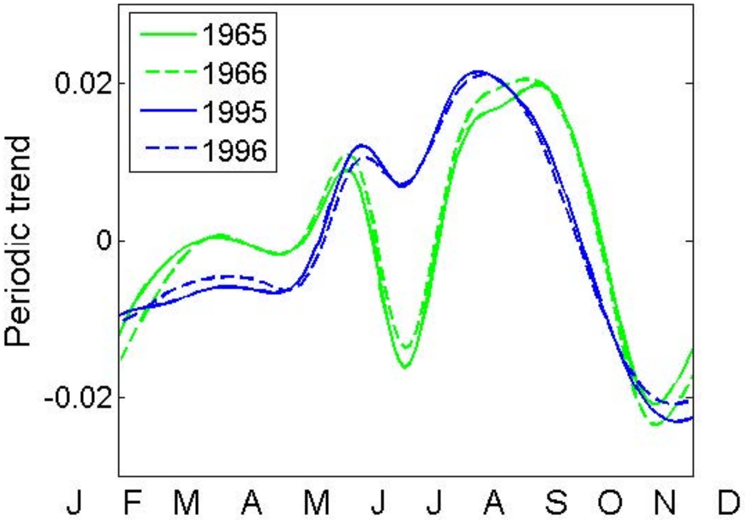
Nondurable goods index 1952-2000



Trends in nondurable goods index

- Long term trend: linear after logarithmic scaling.
- Shocks like the Vietnam War are on a shorter scale, and can result in long-term changes in trend.
- Seasonal trend is complex and evolving over the years.

Four seasonal trend examples



Seasonal trends in nondurable goods index

- Three cycles evident in most years, separated by the Easter/Passover, summer school, and Christmas holidays, respectively.
- Seasonal trends are stable over a couple of years, but evolve over a longer time span.
- The large autumn cycle shows a phase shift between the 60's and 90's, but there is little change in amplitude.
- The small winter cycle is much smaller in the 90's, but the dip due to the summer holidays is much more profound in the 60's.

A concurrent functional linear model approach

We can model the nonseasonal trend plus an evolving seasonal trend as follows:

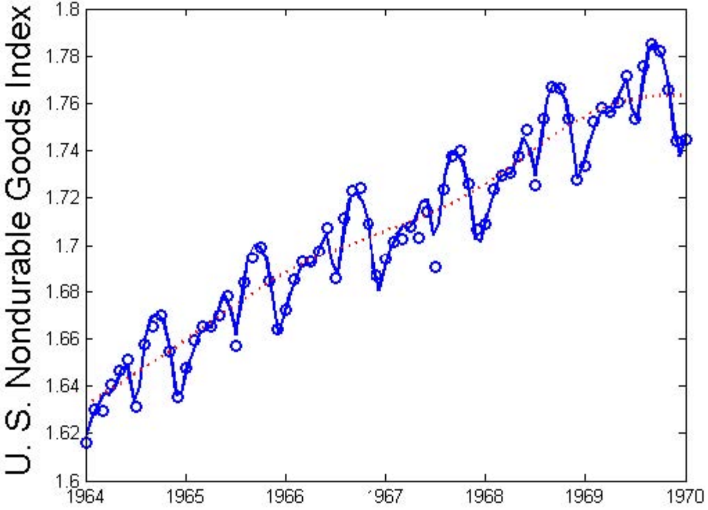
$$\begin{aligned} y(t) &= \alpha(t) + \beta_1(t)\sin(2\pi t/365) + \beta_2(t)\cos(2\pi t/365) + \dots \\ &+ \beta_{2p-1}(t)\sin(2p\pi t/365) + \beta_{2p}(t)\cos(2p\pi t/365) + \epsilon(t). \end{aligned}$$

- Intercept function $\alpha(t)$ captures nonseasonal trend.
- Seasonal trends are represented by the sinusoidal functions with different periods.
- $p = 5$ in our analysis which essentially includes seasonal trends of five different frequencies.

Fitting details

- Analysis conducted on the whole long time series that contains 577 monthly values in the years 1952 to 2000.
- Intercept function α was modeled by B-splines with knots at each year and regularized with $\lambda = 0.01$.
- Each regression function β_j had 7 B-spline basis functions.
- A total of 121 parameters were estimated from 577 data points.

The fit over years 1964-1970



Concurrent functional linear models

- The concurrent functional linear model offers a simple way of relating a functional response to functional covariates.
- However, the influence is simultaneous, and does not permit a covariate to affect the outcome at any time other than the present.
- The model can also be fit to a single long time series provided that the number of parameters is kept small and/or regularization is used.